UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Language proficiency in native and nonnative speakers: an agenda for research and suggestions for second-language assessment

Hulstijn, J.H.

DOI
10.1080/15434303.2011.565844
Publication date
2011
Document Version
Final published version
Published in
Language Assessment Quarterly

Link to publication

Citation for published version (APA):
Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: an agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly*, *8*(3), 229-249. https://doi.org/10.1080/15434303.2011.565844

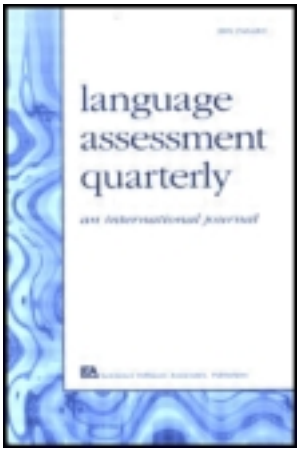
General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Oca 2022

This article was downloaded by: [UVA Univeriteitsbibliotheek]
On: 29 August 2011, At: 01:26
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language Assessment Quarterly
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hlaq20

Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second-Language Assessment

Jan H. Hulstijn [a]
[a] University of Amsterdam

Available online: 16 Aug 2011

To cite this article: Jan H. Hulstijn (2011): Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second-Language Assessment, Language Assessment Quarterly, 8:3, 229-249

To link to this article: http://dx.doi.org/10.1080/15434303.2011.565844

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Language Assessment Quarterly, 8: 229–249, 2011
Copyright © Taylor & Francis Group, LLC
ISSN: 1543-4303 print / 1543-4311 online
DOI: 10.1080/15434303.2011.565844

Routledge
Taylor & Francis Group

Language Proficiency in Native and Nonnative Speakers: An Agenda for Research and Suggestions for Second-Language Assessment

Jan H. Hulstijn

University of Amsterdam

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

This article addresses the question of what language proficiency (LP) is, both theoretically and empirically. It does so by making a distinction, on one hand, between *basic* and *higher language cognition* and, on the other hand, between *core* and *peripheral components* of LP. The article furthermore critically examines the notion of *level* in most second language (L2) assessment scales, showing that it is confounded with people's intellectual functioning because higher levels of LP cannot be attained by people with lower intellectual, educational, occupational, or leisure-time profiles. It is probably for this reason that the Common European Framework of Reference for Languages (Council of Europe, 2001) fails to consistently distinguish between L2 development and L2 proficiency. The LP construct presented in this article can account for the fact that L2 learners with higher intellectual, educational, occupational, or leisure-time profiles may perform, at a given point in time, both better (i.e., in the domain of higher language cognition) and more poorly (i.e., in the domain of basic language cognition) than native speakers with lower profiles. While offering a research agenda for investigating individual differences in first language and L2 acquisition, the article also presents several implications for L2 assessment.

INTRODUCTION

This article offers an attempt to define the construct of language proficiency (LP) beyond a general statement, such as "a person's overall competence and ability to perform in L2 [second language]" (Thomas, 1994, p. 330, footnote 1). Defining a construct is no trivial matter. Formulating a definition is part of the business of constructing a theory that aims to explain puzzling phenomena (Popper, 1959). Defining LP is part of explaining individual differences in the attainment of a first language (L1) or a second language (L2). The article proposes definitions of L1 proficiency (L1P) and L2 proficiency (L2P), distinguishing between basic and higher language cognition. The purpose of the distinction is to define an agenda for research on LP in adult native speakers (L1-ers) and adult nonnative speakers (L2-ers). This agenda pertains both to some truly fundamental issues in the understanding of individual differences in monolingual and bilingual people and to matters of L2P assessment.

Correspondence should be sent to Jan H. Hulstijn, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, The Netherlands. E-mail: j.h.hulstijn@uva.nl

The article has a rather straightforward structure. The LP in Adult Native Speakers and LP in Adult Nonnative Speakers sections deal with L1P and L2P issues, respectively. They are rather long sections with several subsections, using headings to make the article's structure visible. The Implications for L2 Assessment section offers several potential implications for language assessment, whereas the short final section presents some concluding remarks.[1]

LANGUAGE PROFICIENCY (LP) IN ADULT NATIVE SPEAKERS

Theories of L1 and L2 acquisition have to explain the following puzzling phenomena. First, all L1-ers can communicate successfully with each other to a certain degree; they appear to share a basic knowledge of their language. At the same time, however, they differ enormously in the command of their L1. Even a brother and a sister can diverge extensively in the command of their L1, in listening, speaking, reading, and writing skills. Second, L2-ers differ in the success with which they acquire their L2. Even if they are exposed to the same L2 input over the same course of time and spend the same time and effort in trying to learn the language, they will not all end up attaining the same level of L2P. How can we explain these individual commonalities and differences among L1-ers and L2-ers? In this section I focus on L1, proposing the constructs of basic and higher language cognition, which might help explain commonalities and differences in L1P. The LP in Adult Nonnative Speakers section is concerned with L2P.

Basic and Higher Language Cognition

To explain commonalities and differences in language skills among adult native speakers, I postulate two kinds of language ability: basic language cognition (BLC) and higher language cognition (HLC). BLC is what all native speakers have in common; HLC is the domain where differences between native speakers can be observed.

BLC pertains to (a) the largely implicit, unconscious knowledge in the domains of phonetics, prosody, phonology, morphology and syntax; (b) the largely explicit, conscious knowledge in the lexical domain (form-meaning mappings), *in combination with* (c) the automaticity with which these types of knowledge can be processed. BLC is restricted to frequent lexical items and frequent grammatical structures, that is, to lexical items and morphosyntactic structures that may occur in any communicative situation, common to all adult L1-ers, regardless of age, literacy, or educational level.[2] The speed with which linguistic information can be processed may change across the life span, increasing first and decreasing later, from the age of around 30 into old age.

[1]The present article presents some of the main ideas of a book project and, because of its limited length, can only scratch the surface of its topics. As the book project, under the preliminary title *Language Proficiency*, is still in progress, comments on the ideas presented in this article will be highly appreciated.

[2]In Hulstijn (2007) I called this construct *core language proficiency*, but that name appeared to evoke unintended connotations among fellow researchers. For instance, for some people the word *proficiency* implies interindividual variability along one or several scales, whereas the essence of the construct is that there is hardly any variability. I have therefore changed the construct's name to *basic language cognition* (BLC).

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

BLC's processing dimension reflects the fact that speaking and understanding speech (as well as reading and writing) involve parallel processing of phonetic-phonological, lexical, and grammatical information in high speed. Adult native speakers with normal hearing can speak and understand by a rate of two to three words per second (Levelt, 1989, p. 22). Parallel processing, however, comes at the price of momentary disfluences (pauses, repeats, self-corrections) and even uncorrected errors in speech (articulation, lexis, or grammar). Some native speakers are more affected by disfluences and errors in their speech than others (e.g., Goldman-Eisler, 1968).

BLC is restricted to speech reception and speech production; it does not comprise reading and writing. In line with a tradition in linguistics going back to American and European structuralism (Bloomfield, 1933; De Saussure, 1916), I regard the reception and production of speech as a more fundamental human attribute than literacy skills.

HLC is the complement or extension of BLC.[3] HLC is identical to BLC, except that (a) in HLC, utterances that can be understood or produced contain low-frequency lexical items or uncommon morphosyntactic structures, and (b) HLC utterances pertain to written as well as spoken language. In other words, HLC utterances are lexically and grammatically more complex (and often longer) than BLC utterances and they need not be spoken. HLC discourse pertains to topics other than simple everyday matters, that is, topics addressed in school and colleges, on the work floor, and in leisure-time activities. Whereas in the field of L2 teaching and testing, it is common to speak of levels of L2P (see the section A Juxtaposition of L1P and L2P), it is uncommon that linguists speak of levels of L1P. This is not even common in the diagnosis and treatment of language impairments. Yet one could say that there are levels of HLC if one wanted to argue that higher levels of education imply higher levels of HLC. This would mean that talking of levels of L1P would only make sense in the domain of HLC, not in the domain of BLC. I come back to this point in the Language Proficiency in Adult Nonnative Speakers section.[4]

Hypotheses

The distinction between BLC and HLC leads to the following hypotheses.

H1: All adult L1-ers (not suffering from mental disorders), regardless of differences in age and intellectual functioning, are able to comprehend and produce, both correctly and quickly, isolated utterances consisting of high-frequency lexical phrases and high-frequency morphosyntactic structures, when these utterances are perceived under normal acoustical conditions. The null hypothesis in the lexical and grammatical domains, then, is that there is no lexical item or grammatical structure that all native speakers have knowledge of and can process correctly, receptively or productively. In other words, the null hypothesis is that there is no vocabulary or grammar, shared by all native speakers.

H2: Individual differences among adult L1-ers will be relatively large in tasks involving HLC discourse, in all four modes of language use (reading, writing, listening, and

[3] An alternative label of HLC is *extended language cognition*.

[4] One anonymous reviewer rightly pointed out that "obviously deaf populations that use sign languages develop a form of BLC, and perhaps often also HLC, that do not involve speech, and many deaf people fluent in a sign language can communicate effectively through reading and writing as well." This issue, which is beyond the scope of the present article, will be addressed in the book mentioned in footnote 1.

speaking) but almost all adult L1-ers will perform at ceiling in BLC tasks, that is, conceptually simple oral tasks (listening and speaking) involving highly frequent linguistic units.

H3: Although the speed with which humans can process information increases over time until it reaches a peak around the age of 22, and from the age of 27 on gradually decreases (Salthouse, 2009), the vast majority of older people remain capable of processing linguistic information fast enough to allow for relatively unimpaired functional language use, provided that they continue to practice their language skills on a daily basis and do not suffer from severe mental disorders. This claim holds for all languages someone has acquired. In other words, continued language use modulates the decline in processing speed in old age, whereas the continued use of written language is modulated by level of education, type of occupation, and interests.

The definition of BLC and HLC immediately raises the question of where exactly each individual morphological and syntactic structure and each word or expression must be located: in BLC (shared by all L1-ers) or in HLC? Hence, one can wonder whether Hypotheses 1 and 2 are falsifiable. They are, however, if we simply begin with some operationalization of the frequency concept, using, for instance, corpora of oral and written language. We should not expect to find a rock-solid border line between BLC and HLC. Instead, the distinction should be found in terms of prototypicality. The claim is that differences between L1-ers with high and low educational-occupational profiles (in terms of the accuracy and speed with which they perform oral and written tasks) are substantially larger when participants process utterances containing less common linguistic elements than when they process utterances containing more common elements.

Theories and hypotheses are to be seen not in the first place as attempts to render the truth but as strategic tools fuelling empirical research, likely to increase our understanding of a fundamental issue. Thus, by trying to falsify the hypotheses just mentioned (the sooner the better), we will increase our understanding of the commonalities and differences between native speakers in terms of (a) their language knowledge and (b) the use of their language knowledge more or less quickly and fluently, performing various listening, speaking, reading, or writing tasks. Since Chomsky (1965) claimed that all adult native speakers share the same grammatical competence ("the ideal speaker-hearer's intrinsic competence"; p. 4), most researchers, except sociolinguists and speech therapists, have simply taken the proposition for granted, neglecting the obligation of finding out to what extent it can be empirically upheld.

Basic Language Cognition (BLC) and Higher Language Cognition (HLC) Versus Basic Interpersonal Communicative Skills (BICS) and Cognitive Academic Language Proficiency (CALP)

The notions of basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP), proposed by Cummins (1980a, 1980b), come close to my constructs of BLC and HLC, respectively, but there are important differences in specificity and theoretical context. To my knowledge, Cummins never gave a detailed definition of BICS, probably because of his focus on literacy skills (captured by CALP). In Cummins (1980a, p. 84), BICS is characterized as follows:

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

> However, it is clear that not all aspects of language proficiency are related to cognitive and literacy skills. For example, with the exception of severely retarded and autistic children, everybody acquires basic interpersonal communicative skills (BICS) in a first language, regardless of IQ or academic aptitude. As Chomsky (1965) has pointed out, the phonological, syntactical, and lexical skills necessary to function in everyday interpersonal contexts are universal across native speakers. There are individual differences in the ways in which native speakers manifest these linguistic skills in interpersonal communicative contexts, e.g. oral fluency, but for the most part these differences are not strongly related to cognitive or academic performance.[5]

My definition of BLC is much more specific than Cummins's definition of BICS, in that it explicitly refers to the distinction between language reception and production, to the distinction between representation and online processing of linguistic information, and to particular linguistic domains (phonetics, prosody, phonology, morphology, syntax, and the lexicon). In contrast to Cummins, I doubt whether all people (not affected by severe mental disorders) acquire BICS/BLC irrespective of differences in intelligence or academic aptitude. It is precisely for that reason that I hypothesize that BLC exists without exception for all people and for all linguistic domains. Hypotheses are proposed to be falsified. Empirical research will have to show to what extent individual differences in cognitive abilities such as knowledge of the world (of which vocabulary size forms a rough index), attention-allocation and decision-making abilities, reasoning and problem-solving abilities, and working-memory capacity are associated with individual differences in performing HCL and BLC tasks.

Most crucially, however, there is a difference in purpose between Cummins's notions and mine. Although Cummins launched the construct of BICS to underline the importance of CALP in students' educational success, I postulate the constructs of BLC and HLC as falsifiable attempts to explain commonalities and differences in language ability between people, as mediated by other cognitive abilities. In other words, although Cummins's notions of Common Underlying Proficiency, BICS, and CALP and his threshold hypothesis were proposed to help solve a practical issue, my notions of BLC and HLC and the accompanying hypotheses aim to help explain a more fundamental problem of understanding individual differences in language ability.

Explaining Commonalities and Differences Among Adult Native Speakers (L1-ers)

If empirical research would provide evidence for a (nontrivial) notion of BLC, the question arises how to explain the fact that all native speakers, not seriously affected by mental disabilities, are able to acquire BLC, whereas they differ so much in the acquisition of HLC.

The issue of how to explain the universal success in the acquisition of BLC pertains to what, in linguistics, is called the learnability issue (also called the poverty-of-the-stimulus argument). The learnability problem can be stated in the following way. A language has a finite number of words. With these words one can make both an infinite number of grammatical sentences and an infinite number of ungrammatical sentences. A grammar is a system that generates (accounts for) precisely the infinite set of grammatical sentences while excluding precisely the infinite

[5] In Cummins (1980b), BICS is not formerly defined but only characterized with three examples: "Basic interpersonal communicative skills (BICS) in L1 *such as accent, oral fluency, and sociolinguistic competence* [emphasis added] may be independent of CALP" (p. 177).

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

set of ungrammatical sentences. The child is exposed to a finite number of sentences (input). How is the child able to acquire (induce) the grammar of the language from this finite input? The learnability literature, going back to Chomsky (1965, pp. 25–26), is too vast to present here in a comprehensive form. Although the debate was dominated by linguists until the mid-1990s, scholars working in the areas of neuroscience and evolutionary biology have entered the debate since then. Currently, a rationalist and an empiricist school of thought can be discerned. Both schools acknowledge that language acquisition is a matter of both nature and nurture, but they attribute different weights to these two groups of causal factors. The rationalist school is mainly represented by linguists who emphasize the crucial role of a language-specific faculty of mind, called Universal Grammar, a blueprint of human languages, with which all humans are claimed to be born. In recent publications, Chomsky has narrowed down Universal Grammar to the principle of recursivity and the operation of Merge, speculating about the genetic basis of this principle (Chomsky, 2007; Fitch, Hauser, & Chomsky, 2005). Linguists and cognitive neuroscientists in the opposite school, usually called the emergentist or usage-based school, argue that although grammar is arbitrary, it is not a self-contained, autonomous system as the generativist claim. Language acquisition can be accounted for in terms of interactions between biological and environmental processes. There is no need to postulate a language-specific faculty of mind; language knowledge is acquired with domain-general mechanisms (e.g., Bybee, 2006; Croft, 1995; Evans & Levinson, 2009; MacWhinney, 1999, 2008; Tomasello, 2008).

The explanation of native speakers' differences in HLC should be found in cognitive abilities in nonverbal domains (e.g., executive control, reasoning and problem-solving abilities, working-memory capacity) in combination with environmental factors, such as exposure to oral and written language at home, in school, and elsewhere (e.g., Carroll, 1993, Chapters 16 and 17). As a result, adults differ widely in literacy, level of education, professional careers, and cultural profiles (Figure 1).

The fundamental question of why almost all people appear to possess the cognitive abilities to succeed in acquiring their L1 to an impressive extent (see the prior definition of BLC) and why people nevertheless differ in intellectual skills, causing substantial differences in L1P (HLC), is likely to remain a mystery for a considerable time, requiring a multidisciplinary approach.

Empirical Evidence for Individual Differences in Adult Native Speakers (L1-ers)

The LP construct proposed here, distinguishing BLC and HLC in adult L1-ers, has not yet been empirically tested, and Cummins's distinction of BICS and CALP, previously mentioned, has hardly been tested either (Diaz, 1985). There are surprisingly few studies that have looked into individual differences among adult native speakers as a function of their age and level of education or socioeconomic status. Studies conducted by linguists targeted people speaking different sociolects or people diagnosed with various language-related impairments. One wonders whether Chomsky's (1965) claim that all native speakers share the same tacit knowledge of their language, widely accepted for many years, may have drawn the attention of linguists away from investigating this claim. The studies summarized in this subsection were mainly conducted by psychologists investigating matters of aging.

As far as the factor of age is concerned, there is abundant evidence of older adults processing verbal information more slowly than younger adults (e.g., Burke, MacKay, & James, 2000; Schneider, Daneman, & Pichora-Fuller, 2002; Schrauf, 2008; Tun & Wingfield, 1999; Waters

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

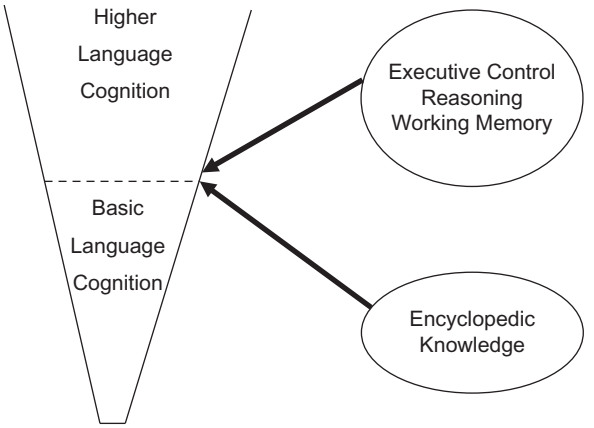


FIGURE 1 Basic and higher cognition. *Note.* The borderline between the two forms of language cognition is represented by a dotted line, to indicate its hypothetical, unattested nature. There is no lid on the cone, to indicate the impossibility of defining maximal language cognition. The arrows illustrate that both BLC and HLC are contingent upon other cognitive abilities (although in different ways and degrees). The cognitive abilities shown in ovals should not be seen as exhaustive.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

& Caplan, 2005). According to Schrauf (2008), it is in tasks involving effortful rather than automatic processing that older adults show decrements relative to younger adults. Retrieval of the meaning of a word as a result of reading or hearing the word is an automatic process, in contrast to word finding, which requires an explicit search through the lexicon. More automatic mental processes are thus not as heavily affected by age as are more explicit mental processes. With respect to the mediating role of education, individuals at higher levels of education have been found to perform better than lower educated individuals in word learning tasks (Van der Elst, Van Boxtel, Van Breukelen and Jolles, 2005) and in picture description tasks (Beland, Lecours, Giroux, & Bois, 1993; Hawkins & Bender, 2002; Le Dorze & Bedard, 1998; Mackenzie, 2000). Dabrowska (1997) found that sentence comprehension may differ as a function of level of education, whereas Chipere (2001) and Dabrowska (2010) observed that intuitions concerning the grammaticality of certain sentences differed as a function of level of education. With respect to the combined effect of age and level of education, higher education has been shown to ameliorate age-related cognitive decline (Bosma, Boxtel, Ponds, Houx, & Jolles, 2003; Meijer, De Groot, Van Boxtel, Van Gerven, & Jolles, 2008).

Recently, Mulder and Hulstijn (2011) assessed, in a sample of 98 adult native speakers of Dutch, how their lexical skills and their speaking proficiency varied as a function of their age (18–76) and level of education and profession. Participants were tested on their lexical knowledge, lexical fluency, and lexical memory, and they performed four speaking tasks, differing in genre and formality. Older participants performed better than younger participants in lexical knowledge, but the opposite was found in lexical fluency and memory. Participants with higher levels of education or profession outperformed participants with lower levels of education or profession in lexical knowledge and lexical memory, but no education effect was observed

in lexical fluency. Level of education or profession was also found to affect the communicative adequacy of the responses in the speaking tasks, but no effect of age was found.

In conclusion, there is substantial empirical evidence for the view that amount of reading and writing activities mediates the negative effects of ageing in that people who read and write much are less affected by information processing and memory problems associated to old age than are people who read and write little. For the constructs of BLC and HLC, the findings appear to mean that, whereas individual differences in BLC tasks are mainly restricted to differences in the speed with which linguistic information can be processed (as a function of age), individual differences in HLC tasks are caused, not only by differences in speed of processing but mainly by differences in intellectual skills and amount of reading and writing activities, as reflected by education, occupation, and leisure-time activities.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

LANGUAGE PROFICIENCY (LP) IN ADULT NONNATIVE SPEAKERS

This section is concerned with L2P. I first look at the notion of components of LP in models of L2P and subsequently examine the notion of level of LP, figuring in language assessment. Then, in the third subsection, I define LP in general, covering both L1P and L2P, suggesting in the last subsection various issues for further inquiry, including research on so called uneven proficiency profiles.

Components of Second-Language Proficiency (L2P)

In the field of L2 teaching and testing it is common to speak of both levels and components of L2P. Over the last 50 years, a variety of models of L2P have been proposed (for a brief overview, see Purpura, 2008). Early models of L2P consisted of a two-dimensional grid, with components of linguistic knowledge along one axis (knowledge of lexis, morphology, syntax, and phonology/orthography), crossed with the four language skills (listening, reading, speaking, and writing) along the other axis, as in the models of Lado (1961) and Carroll (1961/1972). In a seminal article published in 1980, Canale and Swain proposed a model of communicative competence, consisting of grammatical, sociolinguistic, and strategic competence. This influential LP model was later extended by Bachman and Palmer (1996, pp. 66–68), who proposed a three-level hierarchical model of "*language ability*," distinguishing organizational language knowledge (grammatical and textual knowledge), pragmatic language knowledge (functional and sociolinguistic knowledge), and a component of strategic competence (metacognitive components and strategies; see Figure 2).

Scholars who had proposed multicomponent models of LP soon discovered that obtaining empirical support for their models turned out to be extremely difficult. For example, in an early study, Bachman and Palmer (1982) attempted to find evidence for the existence of three traits (grammatical, pragmatic, and sociolinguistic competence), adopting a multimethod design. Participants in the study were 116 English as a Second Language students in the United States, differing widely in age and length of residence. They performed in an oral interview, a writing task, a multiple-choice test, and self-rating. Confirmatory factor analyses produced evidence for

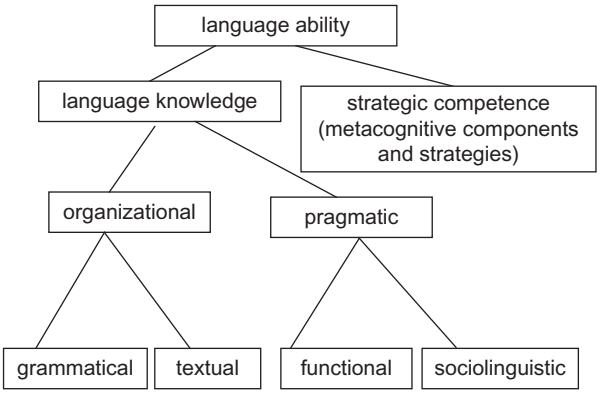


FIGURE 2 Language proficiency model of Bachman and Palmer (1996) (Figure constructed on the basis of Bachman and Palmer [1996, pp. 66–68]).

the existence of a general factor and two specific factors: grammatical/pragmatic competence and sociolinguistic competence.

Harley, Cummins, Swain, and Allen (1990) reported on a study conducted among 175 Grade 6, early French immersion students from the Ottawa region in Canada. The study adopted a three-traits-by-three-methods design. The model components of grammatical, sociolinguistic, and strategic competence formed the three traits, whereas the test methods comprised various oral tasks, written composition tasks, and paper-and-pencil multiple-choice tests. Factor analyses failed to confirm the hypothesized three-trait structure of LP. Instead, a two-factor solution was obtained. The first factor of this solution was interpreted as a "general language proficiency factor" (p. 15). The highest loadings on this general factor came from the three grammar tests. The second factor was interpretable as a "written method factor" (p. 15). Among the reasons suggested for the failure of identifying the three components was the relative homogeneity of the subject sample in terms of French proficiency (Bachman, 1990; Harley, Cummins, Swain, & Allen, 1990).

Sasaki (1993) examined the factor structure of the scores on 11 English L2 subtests, administered to a "homogeneous Japanese EFL student population" (p. 337) consisting of 160 university students, 18 to 23 years old, who had studied English for an average of 7.3 years "through highly controlled formal education" (p. 325). The 11 test scores were derived from three tests: a free composition test, a short-text multiple-choice test (SMC), and a long-text multiple-choice test (LMC), the LMC comprising listening comprehension, reading comprehension, and fill-in-the-gap (cloze). Using covariance structure analysis, Sasaki obtained good fit for a correlated three-component model (composition, SMC, and LMC) as well as for a model in which these three components were not correlated but united in a general second-order factor. No good fit was obtained for a model consisting of one general factor directly associated with all 11 observed variables. This finding goes against Oller's (1983) unitary LP construct. Neither did a model with three independent components (composition, SMC, and LMC) fit the data well. Students also performed three parts of the Japanese version of the Modern Language Aptitude Test (Carroll & Sapon, 1959; paired associates, language analysis, and sound–symbol association) as well as an intelligence test, consisting of 12 subtests, comprising components of verbal intelligence and reasoning. A "reasonable explanation" (p. 335) of the data was provided by a

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

model consisting of a general second-order factor (cognitive ability [CA]) and three first-order factors (aptitude, verbal intelligence, and reasoning). Sasaki then examined the fit of three models, combining LP and CA. No support was observed for a model equating LP and CA; the model taking LP and CA as two separate abilities was not supported either. The correlated model, considering LP and CA as not identical but associated constructs, fitted the data best, with CA explaining 42% of the variance in LP.

Several more recent studies, conducted in the Netherlands, examined the componential structure of reading and writing skills in L1 Dutch and L2 English in 281 high school students (Schoonen et al., 2003; Van Gelderen et al., 2004), and the structure of speaking skills in 181 adult L2 learners of Dutch (De Jong, Steinel, Florijn, Schoonen, & Hulstijn, in press). The design of these studies is similar in that the researchers distinguished between dependent variables (performance in functional reading, writing, or speaking tasks) and independent or predictor tasks. In each study a fairly large number of predictor tasks were included, each representing a potential component of the skill under investigation. A unique feature of these studies is that the battery of predictor tasks includes several tests that measure the speed with which participants can process linguistic information, for example, visual word recognition and sentence verification as predictors of reading, and picture-naming and sentence-building tasks as predictors of writing and speaking. The researchers used analysis of covariance structure, allowing them to test which combination of predictor tests best fit the variance in the dependent variable, while controlling for measurement error to a large extent. A common finding in the reading, writing, and speaking studies is that the variance in the skill under investigation is substantially associated not only with declarative knowledge (vocabulary and grammar) but also with the speed with which linguistic information can be processed.[6]

Hulstijn, Van Gelderen, and Schoonen (2009) suggested that association between performance in the speed tasks and communicative adequacy in the speaking tasks might be explained by the fact that, although knowledge of vocabulary and grammar increases, so does the efficiency with which new knowledge can be processed. "Although *conceptually* skill acquisition can be distinguished from knowledge accumulation, *in reality* knowledge accumulation forms part of skill acquisition because, in real L2 learning, exposure to new words goes hand in hand with exposure to words encountered before" (Hulstijn et al., 2009, p. 576). A main finding in the reading and writing studies was that metacognitive knowledge was observed to be an important predictor. I consider these findings as support for a model of LP, whose "core components" consist of linguistic knowledge and the speed with which this knowledge can be processed (see Figure 3).

Canale and Swain (1980), Bachman and Palmer (1996), and many others were right in including other than purely linguistic competences in the construct of LP (or "language ability," as Bachman and Palmer call it). However, these competences ought to be put in LP's periphery, for the simple reason that although strategic and metalinguistic competences cannot exist without linguistic competences, linguistic competences may exist without strategic and metalinguistic competences. This is supported by the fact that linguistic knowledge came out as the first factor in the factor analyses of the studies conducted by Bachman and Palmer (1982) and by Harley,

[6]For evidence presented in support of the claim that grammatical and pragmatic knowledge form two separate components of LP, see Grabowski (2009). For a review of investigations into the componential nature of the construct of speaking proficiency, see De Jong et al. (in press).

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Language Proficiency	
CORE	PERIPHERY
Linguistic cognition (knowledge and speed) in the phonetic-phonological, morphonological, morphosyntactic, and lexical domains: * Basic language cognition * Higher language cognition	Metacognitive competences * Metalinguistic knowledge * Knowledge of various types of oral and written discourse with their own characteristics * Strategic competences * …

FIGURE 3 Core and peripheral components of language proficiency. *Note.* The distinction between core and periphery is based on the claim that performance in (most) oral and written language tasks is contingent, to a large extent, on more purely linguistic competences and, to a lesser extent, on less purely linguistic competences, such as metacognitive and strategic competences.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Cummins, Swain, and Allen (1990), as well as by the fact that vocabulary knowledge and pronunciation alone explained 75% of the variance in speaking in our study (De Jong et al., in press, Table 2).[7]

Levels of Language Proficiency (LP)

The notion of levels of LP has always been more or less visible in L2 curricula and in the certification of L2P. In L2 instruction, it is common practice to offer courses to learners at beginning, intermediate, and advanced levels. In L2 assessment, all major certificates of speaking proficiency in L2 English and the Oral Proficiency Interview of the American Council on the Teaching of Foreign Languages in the United States distinguish at least four proficiency levels. In this subsection, I examine the notion of level in the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2001), a framework that has become extremely influential in Europe and beyond.

The CEFR finds it roots in the communicative movement (e.g., Wilkins, 1976), and its scales originated from scales developed and used in the United States, Canada, Australia, the United Kingdom, and continental Europe. According to the authors of the CEFR, acquiring an L2 is a matter of development along a horizontal and a vertical dimension (Council of Europe, 2001, pp. 16–17). The *horizontal dimension* consists of two poles. The first pole is concerned with the "language activities" in which language users engage (p. 57) in terms of (a) context of language use, (b) communication themes, and (c) communicative tasks and purposes (pp. 44–56). Chapter 4 of the CEFR contains 40 scales specifying a large number of forms of oral and written language use (pp. 58–84), including several scales of strategic competence. With respect to these strategic scales, one of the CEFR authors, North (2007, p. 656), explicitly refers to the strategic

[7]In the reading and writing studies (Schoonen et al., 2003; Van Gelderen et al., 2004), knowledge of reading and writing strategies were observed to be strong predictors, illustrating the fact that reading and writing typically belong to higher language cognition. See Hulstijn (2011) for a discussion of this finding.

component of the Bachman and Palmer (1996) model. The second pole is concerned with "communicative language competences" (p. 108), subdivided—as in the model of Canale and Swain (1980)—in linguistic, sociolinguistic, and pragmatic competences, apart from various kinds of nonverbal competencies (general knowledge, skills and know-how, existential competence, and ability to learn). Chapter 5 contains 13 scales for the linguistic, sociolinguistic, and pragmatic competences (pp. 110–129).

The *vertical dimension*, applied to all scales, pertains to "an ascending series of common reference levels for describing learner proficiency" (p. 16). In addition, the levels are associated to the notion of L2 development: "Learning which takes place over a period of time needs to be organized into units which take account of progression and can provide continuity" (p. 17). Thus, the levels in the CEFR are associated with both proficiency and development. It is crucial, however, to keep proficiency and development apart. A close examination of the definitions of the B2, C1, and C2 levels in the activity and the competence scales reveals that performance at these higher levels requires higher *intellectual skills* (see Figure 4).

For example, the higher levels of the CEFR's first activity scale, Overall Oral Production, describe language use requiring higher intellectual skills, as can be gathered from the following descriptions: "Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail" (B2), "Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion" (C1), and "Can produce clear, smoothly flowing well-structured speech with an effective logical structure which helps the recipient to notice and remember significant points" (C2). Similarly, the higher levels of most of the competence scales also will normally be attainable only by people with higher levels of education or functioning in higher professions, as is illustrated in the following descriptions of the C2 level: "Has a good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning" (Vocabulary

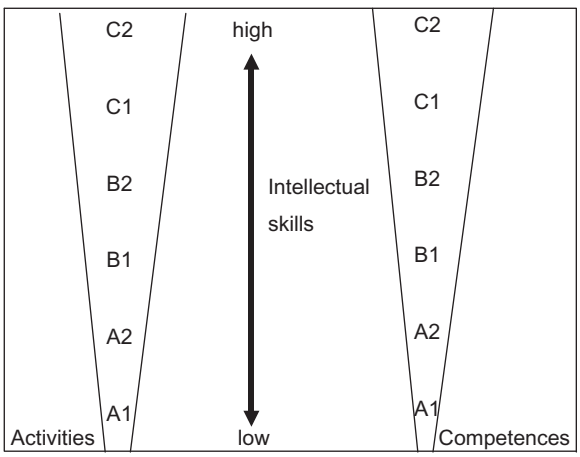


FIGURE 4 Interpretation of the Common European Framework of Reference for Languages's dimensions.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Range, p. 112), "Consistently correct and appropriate use of vocabulary" (Vocabulary Control, p. 112), "Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions)" (Grammatical Accuracy, p. 114), "Writing is orthographically free of error" (Orthographic Control), "Shows great flexibility reformulating ideas in differing linguistic forms to give emphasis, to differentiate according to the situation, interlocutor, etc. and to eliminate ambiguity" (Flexibility, p. 125), "Can create coherent and cohesive text making full and appropriate use of a variety of organizational patterns and a wide range of cohesive devices" (Coherence and Cohension, p. 125), "Can express him/herself at length with a natural, effortless, unhesitating flow. Pauses only to reflect on precisely the right words to express his/her thoughts or to find an appropriate example or explanation" (Spoken Fluency, p. 129).

Two other well-known assessment systems, introduced almost simultaneously with (the pilot version of) the CEFR, namely, DIALANG and the scales of the Association of Language Testers in Europe, likewise show that performance at higher L2P levels requires higher intellectual skills. For example, writing skill at the fifth and highest Association of Language Testers in Europe level is defined as "Can write letters on any subject and full notes on meetings and seminars with good expression and accuracy" (Council of Europe, 2001, p. 251, Appendix D of the CEFR). DIALANG, a self-assessment system developed under the auspices of the Council of Europe, using the same six level labels as the CEFR (Alderson, 2005), defines reading at the C2 level as "I can understand and interpret practically all forms of written language including abstract, structurally complex, or highly colloquial literary and non-literary writings" (Council of Europe, 2001, p. 231, Appendix C of the CEFR).

An exception to this pattern forms the CEFR's pronunciation scale (Phonological Control, p. 117), of which the B2 level is defined as "Has acquired a clear, natural, pronunciation and intonation" and the combined C1 and C2 levels as "Can vary intonation and place sentence stress correctly in order to express finer shades of meaning." To attain this level of phonological control may require a high motivation and a good ear but not necessarily a higher education.

It is important to note that the CEFR authors explicitly acknowledge that the C2 level "is not intended to imply native-speaker or near native-speaker competence" (p. 36). In fact, many adult native speakers will never attain the highest CEFR levels (C1 and C2). However, the authors do not explain why this is so, nor do they explicitly acknowledge that the C1 and C2 levels will generally not be attainable by L2 users with educational backgrounds other than higher education. In other words, they fail to consistently distinguish between L2 development and L2P.

In conclusion, although the CEFR speaks of development and learning with respect to its levels, it is crucial to understand that this should not be taken to mean that every L2 learner, given sufficient learning time and effort, would eventually reach the B2, C1, and C2 levels of the scales in Chapters 4 and 5. The way in which the higher levels are defined, clearly shows that the possibility of attaining these proficiency levels is contingent upon learners' intellectual skills, as reflected by their education, occupation, and leisure-time activities. Any association between CEFR levels of L2P and L2 development as studied in the second language acquisition (SLA) literature would be completely misplaced (Hulstijn, 2010), unless empirical studies show evidence in its support.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

A Juxtaposition of First-Language Proficiency (L1P) and Second-Language Proficiency (L2P)

In the LP in Adult Native Speakers section, I suggested to distinguish between the notions of BLC and HLC for the characterization of L1 proficiency. Although BLC is shared by all adult L1-ers, regardless of age and intellectual skills, HLC reflects people's differences in reading and writing activities. L2P can also be conceptualized in terms of BLC and HLC. My definition of LP in general, covering both L1P and L2P, reads as follows:

> *Language proficiency* (LP) is the extent to which an individual possesses the linguistic cognition necessary to function in a given communicative situation, in a given modality (listening, speaking, reading, or writing). Linguistic cognition is the combination of the representation of linguistic information (knowledge of form-meaning mappings) and the ease with which linguistic information can be processed (skill). Form-meaning mappings pertain to both the literal and pragmatic meanings of forms (in decontextualized and socially-situated language use, respectively). Linguistic cognition in the phonetic-phonological, morphonological, morphosyntactic, and lexical domains forms the center of LP (*core components*). LP may comprise *peripheral components* of a less-linguistic or non-linguistic nature, such as strategic or metacognitive abilities related to performing listening, speaking, reading or writing tasks.
>
> Within LP a base can be defined, called *basic language cognition* (BLC), and an extension, called *higher language cognition* (HLC) (as defined earlier). Whereas all adult native speakers are claimed to share BLC and therefore perform at ceiling in BLC tasks, they are claimed to differ in their HLC profiles (depending on their intellectual skills, education, professional careers and leisure-time activities) and therefore to vary widely in their performance of HLC tasks. Thus, for most adult L1-ers, BLC is only part of their LP. They differ primarily in the realm of HLC. (A possible exception concerns speed of processing, a feature of BLC and HLC, which may show individual differences as a result of aging or other causes, such as hearing impairment.) While L2 learners can acquire HLC in their L2 as native speakers can (depending on their intellectual skills, education, professional careers and leisure-time activities), it remains an open question to what extent postpuberty L2 learners can fully acquire BLC in their L2.
>
> While the notion of *level of proficiency* makes sense in an educational context (curriculum design and language assessment), it has no function in a theory aiming at explaining individual differences in the attainment of a first or second language. Thus, whereas the notions of BLC and HLC are essential in a theory of individual differences, levels of proficiency are defined for practical purposes and the number of levels on L2 proficiency scales is determined on educational rather than principled grounds (cf. CEFR, Ch. 3.5).

As this definition says, it remains an open question to what extent postpuberty L2 learners can acquire BLC in their L2. The SLA literature shows abundant evidence of persistent L1 interference, especially in the domains of auditory and articulatory phonetics. BLC, although being attainable in the domains of vocabulary and most grammatical structures, will generally not be attainable in the domains of pronunciation or with respect to the production of some grammatical features in spontaneous, unmonitored speech (e.g., case assignment and grammatical noun gender in languages that overtly mark these features; e.g., Abrahamsson & Hyltenstam, 2009; Birdsong 2005; Bongaerts, van Summeren, Planken, & Schils, 1997; Van Boxtel, Bongaerts, & Coppen, 2005). However, L2 learners can be as proficient in HLC as L1-ers of the same intellectual, educational, professional, and cultural profile, despite some deficiencies in their L2 BLC. Thus, because "the native speaker" exists only in terms of BLC, the question of whether an L2 learner has attained a native or near-native level of L2P only makes sense with respect to BLC.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Not all studies investigating the age question have controlled for L1 and L2 participants' educational level. The L2-ers in the study of Bongaers et al. (1997), who were indistinguishable from L1-ers, all had received a higher education, whereas the participants in the study of Abrahamsson and Hyltenstam (2009) minimally had 12 years of schooling (no less than senior high school). It is important that in future studies investigating the age question, level of education or intellectual functioning of both the native and nonnative participants be controlled. In general, the empirical investigation of this question would require that L2-ers be compared to L1-ers of the same age and the same level of intellectual functioning, because the hypothesis, advanced in this article, that BLC is shared by all adult L1-ers, regardless of intellectual skills and regardless of age (except for the speed of processing component), has not been investigated and, as I argued here earlier, is unlikely to be correct for all of BLC's features.[8]

Research on Uneven Profiles

An additional research area concerns the matter of so called "uneven profiles" (Council of Europe, 2001, p. 17). First, as one of the CEFR authors stated (B. North, personal communication, April 16, 2010), "uneven profiles" (Council of Europe, 2009, p. 43) are the rule and "flat profiles" (Council of Europe, 2009, p. 43) are the exception. If this is so, research is needed into the extent to which profiles can actually be uneven. For example, how linguistically imperfect (in terms of vocabulary, grammar, pronunciation/intonation, articulation speed) can performance on a C1 task be without failing as a communicative act, and to what extent can weaknesses in one component of linguistic competence be compensated with strengths in another component at a given CEFR level? These questions appear to be particularly relevant at higher proficiency levels (B2, C1, and C2). Second, research is needed on how little linguistic competence is minimally required to perform tasks at the lower levels (A1, A2, and B1). Vocabulary appears to be the most important linguistic component at the lower levels. But which grammatical and phonotactic elements must a learner minimally control at these levels in the case of languages typologically as divergent as Chinese, Japanese, Finnish, and English? Note that research on these questions is particularly needed in the productive skills (speaking and writing).

The question of what learners need to know in the domains of listening and reading appears to be somewhat less pertinent because receptive knowledge minimally required can be deduced from the analyses of corpora of language spoken and written in the activities specified in Chapter 4 of the CEFR. However, there as well, it would be relevant to conduct research using real L2 learners to find out which linguistic skills are minimally needed to understand speech and text in the activities specified in the CEFR's Chapter 4.

IMPLICATIONS FOR SECOND LANGUAGE (L2) ASSESSMENT

In this section, I consider in what ways L2 assessment might profit from the LP construct proposed here (defined in the section A Juxtaposition of L1P and L2 P).

[8]A related issue, beyond the scope of this article, concerns associations between L1P and L2P (see Hulstijn, 2011).

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Transparency Concerning Implied Intellectual Level

First I would like to call for maximal transparency with respect to the notion of *level of LP*. The C1 and C2 definitions quoted in the Levels of Language Proficiency section clearly show that tasks at these levels can be successfully completed only by people with fairly high intellectual skills. The examination boards commissioning tests and the testing industry producing tests should provide maximal transparency with respect to the levels of education or intellectual functioning generally implied in exams at the higher levels of L2P (B2, C1, and C2). A good example of transparency in the Netherlands is the existence of two exams for adult Dutch L2 learners, commissioned by the minister of education. Both exams assess whether test takers can be considered independent L2 users (CEFR Level B). One exam, called the State Exam I, is targeted at people with lower educational backgrounds, whereas Exam II is targeted at people with higher levels of education. Thus, all stakeholders are advised *not* to regard Exams I and II as Steps 1 and 2 on an L2P ladder. Candidates are advised to sit for one exam only, not two.

Furthermore, the industry should—much more emphatically than so far has been the case—state (a) that L2 exams are no means to test whether a candidate has a native or near-native command of the L2, (b) that "the" native speaker level of LP does not exist, and (c) that native speakers differ enormously in their language skills, mainly because of differences in age, intellectual skills, education, occupation, and leisure-time activities.

Finally, examination boards and test producing institutions should provide information about the way in which the two poles of the CEFR's horizontal dimension (activities and competences) are tested. In particular, it should be clear with which types of criteria (scales) performance in speaking and writing will be rated. Will this be mainly criteria of communicative adequacy regardless of linguistic accuracy, or will performance also or mainly be rated in terms of linguistic competences? The view on LP exposed in this article is clear in this respect: Research clearly shows that the linguistic competences form the core of communicative competence (Components of L2P section).

Testing of Core Components Separately

Second, I would like to call for a revival of the testing of core components of LP, in particular control of vocabulary and grammar. If we want to assess whether a candidate has attained the C1 level of Overall Oral Production ("Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion"; Council of Europe, 2001, p. 58), a 5-min oral interview in which the candidate is asked to talk about several level-appropriate topics does not suffice. A level-appropriate productive vocabulary test should complement the interview simply because it is impossible to ascertain in a 5-min interview that the candidate has a command of, say, the 15,000 most frequent content words of the language, required when one wants to give "clear, detailed descriptions and presentations on complex subjects," previously quoted. Similarly, for a writing test at the B2 level ("Can write clear, detailed texts on a variety of subjects related to his/her field of interest, synthesizing and evaluating information and arguments from a number of sources"; p. 61), a single writing assignment does not suffice. Ideally, one would need to give candidates at least five assignments. Obviously, this would be rather expensive. However, the

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

validity of the single-text assignment would considerably increase, if separate tests of control of vocabulary, grammar, and spelling at the B2 level of the corresponding scales in Chapter 5, were administered.

Two-Phase Assessment

To increase test efficiency, one could think of a two-phase assessment, involving the administration of both discrete-point and integrative-skill tests (Carroll, 1961/1972). In the first phase, one or several relevant linguistic competences are tested with discrete-point tests, using nonauthentic tasks. Good content validity and high reliability can relatively easily be attained with this type of test. In the second phase, candidates take one or more integrative-skill tests of functional language use, in (quasi-) authentic tasks. The validity of the integrative-skill testing in Phase 2 will increase, when, in Phase 1, it has first been ascertained that candidates have sufficient linguistic knowledge at their disposal. To increase efficiency and to keep the cost of testing low, one could require that only candidates who have passed the Phase 1 discrete-point tests are allowed to sit the Phase 2 integrative-skill test(s). I would like to emphasize that this proposal does *not* mean a return to the assessment practices of the middle of the previous century, because the blessings of the communicative movement come to bear in the second phase of the exam. A comparison with a driver's test might be appropriate here. In many countries, candidates first take a (computer-administered) test of knowledge of traffic rules and regulations. Only if this test has been passed is the candidate allowed to proceed to the much more costly but crucial road test itself. Just like a driving license should not be issued without the candidate having passed the practice test on the road, a L2P certificate should not be issued without an integrated skill test. The two-phase proposal applies only to L2P exams at the B1 (or A2) level and up, not at the A1 (or A2) level, because when candidates only have to know a limited number of expressions for a limited number of speech acts, it does not make sense to test this knowledge in two phases.

Assessment of First-Language Proficiency (L1P)

In several countries, the CEFR has begun to impact on the secondary-school curriculum for the national language, the language that, for many students, is the L1. In the Netherlands, in 2009, a system akin to the CEFR was proposed for the school subject Dutch ("Expertgroep Doorlopende Leerlijnen," 2009). The system consists of four domains (oral interaction, reading, writing, and metalinguistic knowledge) and four levels. The levels match the school types in the lower and higher streams of the secondary-school system. The contents of the CEFR are not literally copied to the L1 curriculum, but the CEFR's skeleton, consisting of a horizontal dimension (in turn consisting of an activity pole and a competence pole) and a vertical dimension of levels, underlies the proposed L1 framework too. The proposed system pertains not to what I call basic language cognition but to what I call higher language cognition, that is, to communicative tasks that students in the lower and higher tracks of the school systems should aim to attain. In principle, this approach need not be met with skepticism. It is good to formulate targets for each subject in each school type and to test their attainment accordingly, as long as it is understood by all stakeholders that the levels of the proposed system match students' intellectual skills and that it thus does not make sense to administer an L1 test of, say, Level 4 to students placed in the Level

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

2 track. Transparency is needed here as it is in L2P assessment (see the Transparency Concerning Implied Intellectual Level section).

Do We Need a Test of Basic Language Cognition in Second Language (L2)?

Consider the case of a fictitious woman—let us call her Sofia—who at age 30 marries a foreigner, a speaker of another language than her L1. Sofia, who has completed only elementary school in the country of origin, joins her new husband, migrates to his country, and finds a job in a grocery shop, soon beginning to pick up the language of this country (her L2) in the home and work environment. In a few years' time, she has picked up so much of her L2 that she can converse with the native speakers in her immediate environment without difficulty. Given the definition of BLC (Basic and Higher Language Cognition section), she may well have acquired a great deal of BLC—probably not all of it in the productive mode, as her speech exhibits features of L1 transfer, but perhaps almost all of it in the receptive mode. She can parse and understand the speech spoken in her environment automatically and without difficulty. Her husband has—by definition—acquired BLC completely because he is an adult native speaker and does not suffer from language-associated disabilities. The question now is, would it make sense to develop tests to assess BLC in L1-ers (children and adults, like Sofia's husband) and in L2-ers (like Sofia)? Yes, it certainly would for purposes of testing the hypotheses presented in this article (see Hypotheses section) and possibly also for purposes of diagnosis. However, I do not see why stakeholders other than researchers and people involved in diagnosis would benefit from knowing to what extent L2-ers have acquired BLC. Assessment practices based on the CEFR or a similar *communicative* framework, involving both purely linguistic competences and skills of applying these competences in quasi real-words tasks, is all what is needed. CEFR-based L2 assessment serves the needs of institutions and individuals who want to know what candidates "Can Do" in terms of relevant real-world tasks. This may well mean that many L2-ers who perform at the C2 level of speaking have *not* completely acquired BLC in terms of knowledge and automaticity of processing. However, what is important for all L2-ers is that tests for listening and speaking at the lower levels (A1, A2, and B1 in the CEFR) should not require candidates to possess intellectual skills, knowledge of the world, memory skills, and the like, beyond what adults with little schooling can be assumed to possess. The same goes for lower level reading and writing tests, with the restriction that the candidates are assumed to have mastered the writing system of their L2. In the case of illiterate candidates or candidates who have learned to read and write in a different writing system, the lower level L2 tests do rightly require knowledge and skills beyond what candidates can do in their L1.

CONCLUDING REMARK

In this article I have proposed a construct of LP, distinguishing, on one hand, BLC and HLC (see Basic and Higher Language Cognition section) and, on the other hand, core and peripheral components (see Components of L2P section). The ideas on which these notions are based are not original. In the reading-processes literature, for example, Perfetti (1999) made a distinction between lower level and higher level information processing. And, as I explained in the section

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

BLC and HLC versus BICS and CALP, Cummins's (1980a, 1980b) distinction between BICS and CALP are akin, although markedly different from my notions of BLC and HLC. Furthermore, my proposal to distinguish core from peripheral components of LP (Components of L2P section) builds upon earlier proficiency models of Canale and Swain (1980) and Bachman and Palmer (1996).

A fruitful hypothesis should be wrong in such a way that the research community feels challenged to prove it wrong. In contrast, a fruitless hypothesis is perceived as either so outright wrong or so plausible that nobody will feel tempted to falsify its claims. I hope that the constructs and hypotheses proposed in this article belong to the fruitful category.

Although my definition of LP is likely to be incorrect in some important respects, this need not invalidate my views on the CEFR (that it does not sufficiently keep development apart from proficiency and that its vertical dimension represents levels of intellectual skills) and the suggestions advanced in the Implications for L2 Assessment section with respect to L2 assessment (transparency, and two-phase testing of core components and integrative skills). I hope that readers will view these suggestions as serious attempts to improve assessment practices, in the interest of all stakeholders, not least in the interest of L2 learners with lower educational backgrounds.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

ACKNOWLEDGMENTS

Work on this article was funded by the Netherlands Organisation for Scientific Research (grant 360-70-230). I thank the four anonymous reviewers and my associates at the University of Amsterdam for their useful feedback on earlier versions of this text.

REFERENCES

Abrahamsson, M., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, *59*, 249–306.

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London, UK: Continuum

Bachman, L.F. (1990). Constructing measures and measuring constructs. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency*. Cambridge (pp. 26–38). Cambridge, UK: Cambridge University Press.

Bachman, L. F., & Palmer, A. S. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, *16*, 449–465.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.

Beland, R., Lecours, A. R., Giroux, F., & Bois, M. (1993). The MT-86 β Aphasia Battery: A subset of normative data in relation to age and level of school education (Part II). *Aphasiology*, *7*, 359–382.

Birdsong, D. (2005). Why not fossilization. In Z.-H. Han & T. Odlin (Eds.), *Studies of fossilization in second language acquisition* (pp. 173–188). Clevedon, England: Multilingual Matters.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, *19*, 447–465.

Bosma, H., Boxtel, M. P. J. van, Ponds, R. W. H. M., Houx, P. J., & Jolles, J. (2003). Education and age-related cognitive decline; the contribution of mental workload. *Educational Gerontology*, *29*, 165–173.

Burke, D. M., MacKay, D. G., & James, L. E. (2000). Theoretical approaches to language and aging. In T. Perfect & E. Maylor (Eds.), *Models of cognitive aging* (pp. 204–237). Oxford, England: Oxford University Press.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, *82*, 711–733.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1–47.

Carroll, J. B. (1972). Fundamental considerations in testing for English language proficiency of foreign students. In H. B. Allen & R. N. Campbell (Eds.), *Teaching English as a second language: A book of readings* (pp. 313–320). New York: McGraw Hill. (Reprinted from a report of the conference *Testing the English proficiency of foreign students,* held in May 1961, published by the Center for Applied Linguistics, Washington, D.C.)

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: Cambridge University Press.

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test manual*. New York, NY: Psychological Corporation.

Chipere, N. (2001). Variations in native speaker competence: Implications for first-language teaching. *Language Awareness*, *10*, 107–124.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (2007). Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies*, *15*, 1–12.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press/Council of Europe.

Council of Europe. (2009). *Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment* (Final version). Strasbourg, France: Author.

Croft, W. (1995). Autonomy and functionalist linguistics. *Language*, *71*, 490–532.

Cummins, J. (1980a). The construct of language proficiency in bilingual education. In J.E. Alatis (Ed.), *Current issues in bilingual education. Georgetown University Round Table on Languages and Linguistics (GURT) 1980* (pp. 81–103). Washington, DC: Georgetown University Press.

Cummins, J. (1980b). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, *14*, 175–187.

Dabrowska, E. (1997). The LAD goes to school: A cautionary tale for nativists. *Linguistics*, *35*, 735–766.

Dabrowska, E. (2010). Naive v. expert intuitions: An empirical study of acceptability Judgments. *The Linguistic Review*, *27*, 1–23.

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (in press). Facets of speaking proficiency. *Studies in Second Language Acquisition*.

De Saussure, F. (1916). *Cours de linguistique générale*. Paris, France: Payot.

Diaz, R. F. (1985). Bilingual cognitive development: Addressing three gaps in current research. *Child Development*, *56*, 1376–1388.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*, 429–448.

Expertgroep doorlopende leerlijnen taal en rekenen [Expert group continuous cirrucula]. (2008). *Over de drempels met taal: De niveaus voor de taalvaardigheid*. Enschede, the Netherlands: Stichting Leerplanontwikkeling.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005. The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*, 179–210.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London, UK: Academic Press.

Grabowski, K.C. (2009). *Investigating the construct validity of a test designed to measure grammatical and pragmatic knowledge in the context of speaking*. Unpublished doctoral dissertation, Teachers College, Columbia University, New York, NY.

Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 7–38). Cambridge, England: Cambridge University Press.

Hawkins, K. A., & Bender, S. (2002). Norms and relationship of the Boston Naming Test performance to vocabulary and education: A review. *Aphasiology*, *16*, 1143–1153.

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, *91*, 663–667.

Hulstijn, J. H. (2010). Linking L2 proficiency to L2 acquisition: Opportunities and challenges of profiling research. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (pp. 233–238). Eurosla Monographs Series, 1. Retrieved from http://eurosla.org/monographs/EM01/EM01home.html

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011

Hulstijn, J. H. (2011). Explanations of associations between L1 and L2 literacy skills. In M.S. Schmid & W. Lowie (Eds.), *Modeling bilingualism: From structure to chaos* (pp. 85–111). Amsterdam, the Netherlands: John Benjamins.

Hulstijn, J. H., Van Gelderen, A., & Schoonen, R. (2009). Automatization in second-language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics*, *30*, 555–582.

Lado, R. (1961), *Language testing*. New York, NY: McGraw-Hill.

Le Dorze, G., & Bedard, C. (1998). Effects of age and education on the lexico-semantic content of connected speech in adults. *Journal of Communication Disorders*, *31*, 53–71.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Mackenzie, C. (2000). Adult spoken discourse: The influences of age and education. *International Journal of Language and Communication Disorders*, *35*, 269–285.

MacWhinney, B. (Ed.). (1999). *The emergence of language*. Mahwah, NJ: Erlbaum.

MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). New York, NY: Routledge.

Meijer, A., de Groot, H. M., van Boxtel, P. J., van Gerven, W. M., & Jolles, J. (2008). Are age differences in verbal learning related to interstimulus interval and education? *Experimental Aging Research*, *34*, 323–339.

Mulder, K., & Hulstijn, J. H. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*. Advance online publication. doi:10.1093/applin/amr016

North, B. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal*, *91*, 656–659.

Oller, J. W., Jr. (1983). Evidence for a general language proficiency factor: An expectancy grammar. In J. W. Oller, Jr. (Ed.), *Issues in language testing research* (pp. 3–10). Rowley, MA: Newbury House.

Perfetti, C. (1999). Comprehending written language: A blueprint for the reader. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 167–208). Oxford, UK: Oxford University Press.

Popper, K. R. (1959). *The logic of scientific discovery*. London, UK: Hutchinson.

Purpura, J. E. (2008). Assessing communicative language ability: Models and their components. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education, Vol. 7. Language testing and assessment* (2nd ed., pp. 53–68). Dordrecht, the Netherlands: Kluwer.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*, 507–514.

Sasaki, M. (1993). Relationships among second language proficiency, foreign language aptitude, and intelligence: A structural equation modeling approach. *Language Learning*, *43*, 313–344.

Schneider, A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, *56*, 139–152.

Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: the role of linguistic fluency, linguistic knowledge and metacognitive knowledge. *Language Learning*, *53*, 165–202.

Schrauf, R.W. (2008). Bilingualism and aging. In J. Altarriba & R. R. Heredia (Eds.), *An introduction to bilingualism: Principles and processes* (pp. 105–127). New York, NY: Erlbaum.

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning*,*44*, 307–336.

Tomasello, M. (2008). *The origins of communication*. Cambridge, MA: MIT Press.

Tun, P. A., & Wingfield, A. (1999). Language and communication: Fundamentals of speech communication and language processing in old age. In A. D. Fisk & W. A. Rogers (Eds.), *Handbook of human factors and the older adult* (pp. 125–149). New York, NY: Academic Press.

Van Boxtel, S., Bongaerts, T., & Coppen, P.-E. (2005), Native-like attainment of dummy subjects in Dutch and the role of the L1. *International Review of Applied Linguistics in Language Teaching*, *43*, 355–380.

Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, 11, 290–302.

Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in first- and second-language reading comprehension: A componential analysis. *Journal of Educational Psychology*, *96*, 19–30.

Waters, G., & Caplan, D. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory*, *13*, 403–413.

Wilkins, D. A. (1976). *Notional syllabuses*. Oxford, UK: Oxford University Press.

Downloaded by [UVA Univeriteitsbibliotheek] at 01:26 29 August 2011